

УДК 517.968, 57

© *М. Н. Назаров*

БАЗОВАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДЛЯ ОПИСАНИЯ ПРОЦЕССОВ РЕГУЛЯЦИИ БИОСИНТЕЗА БЕЛКОВ

В работе предложена модель для описания кинетики основных процессов регуляции биосинтеза белков. При построении искомой модели к этим процессам в первую очередь были отнесены регуляция транскрипции РНК, посттранскрипционная обработка РНК, ядерный транспорт для РНК и белков, а также регуляция трансляции РНК и посттрансляционная обработка белков. Главной задачей данной модели является обеспечение базового инструмента для моделирования внутриклеточной биохимии, который будет опираться только на физически строгие уравнения и биологически обоснованные параметры. Для достижения этой цели мы разделяем переменные на два блока: численности в ядре и численности в цитоплазме клетки для белков и РНК, а также применяем принцип геометрической вероятности при выводе уравнений модели. При этом все воздействия на уровне транскрипции моделируются с помощью специальных переменных — вероятностей присоединения регуляторного комплекса к гену. Частным примером такого комплекса является РНК-полимераза, и, соответственно, изменяя вероятность ее присоединения к гену, осуществляется прямое воздействие на частоту инициации транскрипции.

Ключевые слова: внутриклеточная биохимия, генные сети, регуляция генов, ядерный транспорт.

DOI: [10.20537/vm160406](https://doi.org/10.20537/vm160406)

Введение

Белковые ферменты играют одну из центральных ролей в организации и управлении биохимическими процессами в живых клетках. Поэтому регуляция синтеза белков — это важнейший механизм, который позволяет контролировать специализацию клетки, а также ее рост и деление (см. [1]). Реализовать подобную регуляцию можно на разных этапах биосинтеза белков: транскрипции РНК в ядре клетки, посттранскрипционной обработки РНК в ядре, транспорта РНК из ядра в цитоплазму, трансляции белков на основе РНК в цитоплазме, посттрансляционной обработки белков, а также транспорта белков из цитоплазмы в ядро (подробнее об этих механизмах см. [1]). Основными катализаторами, которые реализуют регуляторные воздействия на данных этапах, являются как сами белки, так и собранные на их основе белковые комплексы. Таким образом, вопрос о регуляции биосинтеза системы белков сводится в первом приближении к вопросу о взаимной регуляции различными белками процессов синтеза друг для друга. В данном контексте широкое распространение получила концепция генных сетей — комплексов из генов, их продуктов (РНК и белков) и взаимосвязей между ними (см. [2–5]). Построение математических моделей для генных сетей является важным направлением в современной биологии, поскольку позволяет получить удобный инструмент для анализа внутриклеточной биохимии. Однако все существующие модели генных сетей в той или иной степени используют один и тот же набор методологических упрощений.

- Игнорируются промежуточные звенья в реализации регуляторных воздействий.
- Основное внимание уделяется выявлению связей и общей топологии генной сети, а не точности описания динамики с изменением концентраций молекул во времени.
- Для моделей чаще всего используются функциональные зависимости, чей вывод не подкреплен строгим физическим обоснованием.

- Коэффициенты, которые используются в уравнениях моделей, часто также не имеют конкретного биологического смысла.

В некоторой степени обозначенные упрощения являются взаимосвязанными. Так, исходная установка на выявление топологии геной сети и пренебрежение точным описанием кинетики приводит к отказу от рассмотрения промежуточных звеньев в регуляторных воздействиях. В свою очередь, эти два пункта приводят к тому, что нет никакой необходимости строить физически строгую модель с формальным выводом и обоснованием всех уравнений, а также снимает ограничение на использование технических параметров в уравнениях, лишенных явного биологического смысла. К сожалению, все это приводит к существенной проблеме с масштабируемостью моделей и делает интеграцию нескольких моделей геной сетей в одну крупную модель сложной задачей. К примеру, предположим, что мы хотим объединить две базы данных, описывающие различные геной сети одного организма, но у которых промежуточные регуляторные агенты взаимосвязаны. В этом случае чистое объединение этих баз данных будет некорректным решением, так как не учтет взаимосвязи между промежуточными регуляторными агентами, которые не включены в явном виде в классические модели.

В рамках данной работы мы предлагаем отойти от традиций классических моделей геной сетей и строить такие модели на основе следующих требований:

- 1) учитывать промежуточные молекулы в реализации регуляторных воздействий;
- 2) уделять особое внимание кинетике всех вовлеченных процессов;
- 3) для всех уравнений модели использовать только физически строгие основания;
- 4) для всех параметров модели потребовать наличия у них явного биологического смысла, а также потенциальной измеримости.

Перед тем как переходить непосредственно к изложению самой модели, мы рассмотрим примеры процессов регуляции биосинтеза белков, а также обсудим, как эти процессы мы будем вводить в итоговой модели.

§ 1. Описание ключевых процессов регуляции биосинтеза белков

Первый и наиболее очевидный способ регуляции биосинтеза белков сводится к изменению частоты транскрипции¹ РНК. Примерами подобных механизмов являются:

- 1) присоединение фактора транскрипции к промотору² гена; в зависимости от типа молекулы такой комплекс на промоторе гена может либо повысить частоту присоединения РНК-полимеразы к стартовому участку кодирующей последовательности гена (активатор), либо понизить эту частоту (репрессор); также может увеличиться или уменьшиться вероятность присоединения к промотору регуляторной молекулы другого типа;
- 2) присоединение энхансера³ к регуляторному комплексу на промоторе гена либо фиксация энхансера в пространстве ядра близко к промотору гена; близость энхансера к промотору гена повышает шанс присоединения факторов транскрипции.
- 3) ремоделирование хроматина — изменение плотности упаковки ДНК; конкретный пример такого механизма — это ацетилирование белков гистонов, на которые наматывается ДНК в ядре; в общем случае при высокой плотности упаковки ДНК синтез РНК с данного гена РНК-полимеразой будет невозможен;

¹Частоты считывания РНК с ДНК гена, который осуществляется с помощью РНК-полимеразы.

²Промотор — это регуляторная часть гена, которая расположена перед кодирующей частью.

³Энхансеры — это небольшие сегменты ДНК, которые способны формировать устойчивые комплексы с факторами транскрипции. Энхансеры могут быть расположены на последовательности ДНК достаточно далеко от регулируемых им генов.

- 4) метилирование ДНК — присоединение метильной группы к цитозину; в подавляющем большинстве случаев метилирование CpG-динуклиотидов в промоторной области гена привлекает факторы ремоделирования хроматина и приводит к полному подавлению транскрипции за счет плотной упаковки ДНК гена.

Замечание 1. В рамках данной модели все воздействия на уровне транскрипции мы будем описывать с помощью специальных переменных $I_{p_k^r}(g_i)$ — вероятностей присоединения регуляторного комплекса p_k^r к гену g_i . Само же регуляторное воздействие комплекса p_j^r на ген g_i мы будем сводить к изменению других вероятностей $I_{p_k^r}(g_i)$, не вдаваясь в подробности: осуществлено ли это за счет энхансеров, факторов транскрипции, ремоделирования хроматина или метилирования ДНК. При этом сам фермент РНК-полимераза, осуществляющая транскрипцию в модели, вводится как p_1^r .

Второй способ регуляции биосинтеза белков — это посттранскрипционное воздействие на РНК. Примерами подобных механизмов являются:

- 1) сплайсинг — вырезание интронов из РНК; в частности, в результате альтернативного сплайсинга можно на основе одной базовой последовательности РНК получить множество разных мРНК и, как следствие, множество разных белков;
- 2) РНК-интерференция — связывания молекулы РНК с РНК-белковым комплексом RISC, содержащим комплементарную для нее последовательность РНК; после связывания эта молекула не может использоваться для синтеза белка, и в большинстве случаев подвергается деградации и разрывается на отдельные основания.

Замечание 2. В модели данный уровень регуляции реализован за счет введения подмодели химической кинетики, отвечающей за реакции в ядре $\text{Reaction}_{r_i}^{(2)}$ и цитоплазме $\text{Reaction}_{r_i}^{(1)}$.

Третий способ регуляции биосинтеза белков сводится к управлению транспортом РНК и белков между ядром и цитоплазмой. Примерами таких механизмов будут:

- 1) диффузионный транспорт молекул малого размера через ядерные поры; ограничен лишь числом пор и базовой скоростью проникновения молекулы через пору; может быть лимитирован в одном направлении и не позволять молекулам перемещаться в обратную сторону через ядерную пору;
- 2) активный импорт молекул через ядерные поры; внутрь ядра импортируются молекулы, содержащие специальную цепочку аминокислот — последовательность NLS (nuclear localization signal); такие молекулы объединяются с белками импортинами, которые и проводят активный импорт молекулы через ядерную пору;
- 3) активный экспорт молекул через ядерные поры; за пределы ядра экспортируются молекулы, содержащие специальную цепочку аминокислот — последовательность NES (nuclear export signal); такие молекулы объединяются с белками экспортинами, которые проводят активный экспорт молекулы через ядерную пору.

Замечание 3. В рамках модели транспорт будет учтен за счет введения двух наборов переменных для РНК, белков и регуляторных комплексов: $N_r^{(1)}$, $N_p^{(1)}$ и $N_{p^r}^{(1)}$ — численности в цитоплазме, $N_r^{(2)}$ и $N_{p^r}^{(2)}$ — численности в ядре. Сам же транспорт мы будем вводить с помощью коэффициентов $v_m^{(2 \rightarrow 1)}$ и $v_m^{(1 \rightarrow 2)}$, задающих средние скорости перехода молекул r , p и p^r между этими двумя областями клетки.

Четвертый способ регуляции биосинтеза белков заключается в непосредственном управлении трансляцией мРНК в белки на рибосомах. Примеры таких механизмов:

- 1) контроль инициации трансляции; наиболее очевидный пример — это полное подавление трансляции за счет блокирования фазы инициации;
- 2) контроль скорости элонгации; управление скоростью синтеза полипептидной цепи возможно за счет регулирования концентраций специфических тРНК в непосредственной близости от рибосом; понижение концентраций определенных тРНК вызовет снижение скорости синтеза белков, в которых они задействованы;
- 3) контроль терминации трансляции; данное регулирующее воздействие может заставить рибосому интерпретировать терминирующий кодон, расположенный в середине цепочки мРНК, как осмысленный; в результате на основе одной мРНК может синтезироваться несколько вариантов белков — удлиненные и укороченные.

Замечание 4. Механизмы контроля трансляции в явном виде не включаются в модель. Фактически осуществление этих механизмов возможно по схожей схеме с регуляцией активности генов за счет вариации параметра $I_{p_2^r}(r_i)$ — вероятности присоединения мРНК r_i к рибосоме и средней скорости трансляции $v_{p_2^r}(r_i)$ мРНК r_i на рибосоме p_2^r .

Пятый способ регуляции биосинтеза белков сводится к посттрансляционной обработке белка. Примерами таких механизмов могут послужить:

- 1) модификация основной цепочки белка; это может быть либо присоединение небольших химических групп к основной цепочке (ацилирование, аргинилирование), либо как вариант расщепления одной или нескольких пептидных связей в белке ферментом протеазой;
- 2) модификация боковых цепочек белка; примером может послужить присоединение гидрофобных групп для локализации белка в мембране, а также присоединение или отщепление небольших химических групп: ацетилирование, метилирование, фосфорилирование, йодирование, гликирование и т. д.

Замечание 5. В модели данный уровень регуляции реализован так же, как и посттранскрипционная обработка — за счет интеграции модели химической кинетики.

§ 2. Базовая модель внутриклеточной биохимии

В рамках модели рассматриваются следующие категории объектов:

- r — молекулы РНК;
- g — гены (их сегменты ДНК);
- p — белки и белковые комплексы, не участвующие напрямую в регуляции генов;
- p^r — белки и белковые комплексы, осуществляющие регуляцию активности генов.

Для определенности транскриптазу будем обозначать p_1^r , а рибосому — p_2^r . Численности для молекул r и p разделяются на две категории:

- 1) $N_r^{(1)}$, $N_p^{(1)}$ и $N_{p^r}^{(1)}$ — численности за пределом ядра клетки;
- 2) $N_r^{(2)}$ и $N_{p^r}^{(2)}$ — численности внутри ядра клетки.

Изменение численности РНК в пределах ядра будет происходить в модели в результате трех процессов: транскрипции, активного и пассивного транспорта между ядром и цитоплазмой,

а также за счет каталитических⁴ реакций. Для определенности положим, что первые n молекул РНК непосредственно считываются с первых n генов клетки.

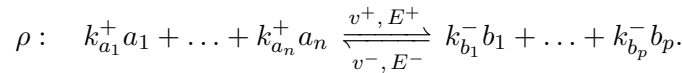
$$\frac{dN_{r_i}^{(2)}}{dt} = v_{p_1^r}(g_i) \left(\frac{V_{p_1^r} N_{p_1^r}^{(2)}}{V_{nuc}} \right) I_{p_1^r}(g_i) + \text{Exchange}_{r_i}^{(1 \leftrightarrow 2)} + \text{Reaction}_{r_i}^{(2)}.$$

В этом выражении параметр $I_{p_1^r}(g_i)$ — это вероятность присоединения транскриптазы p_1^r к стартовому участку гена g_i , константа V_{nuc} — объем ядра, а $v_{p_1^r}$ — средняя скорость транскрипции. В первом приближении мы будем считать скорость $v_{p_1^r}$ постоянной величиной, а в обобщении модели можем сделать ее функцией от концентраций нуклеотидов в ядре, которые требуются для синтеза r_i с гена g_i . Фактор, отвечающий за транспорт r_i , мы разобьем на две части:

$$\text{Exchange}_{r_i}^{(1 \leftrightarrow 2)} = v_{r_i}^{(2 \rightarrow 1)} N_{r_i}^{(2)} - v_{r_i}^{(1 \rightarrow 2)} N_{r_i}^{(1)}.$$

В этом выражении $v^{(1 \rightarrow 2)}(r_i)$ — средняя скорость транспорта молекулы r_i в сторону ядра, $v^{(2 \rightarrow 1)}(r_i)$ — средняя скорость транспорта молекулы r_i в сторону цитоплазмы клетки. В первом приближении мы будем считать эти параметры постоянными, а в обобщении модели можем учесть количество пор в ядре, присоединение последовательностей NES и NLS, а также реализацию транспорта белками импортинами и экспортинами.

В качестве модели химической кинетики для реализации $\text{Reaction}_{r_i}^{(2)}$ можно использовать любую из физически обоснованных и строгих моделей. Например, можно использовать модель из работы [6], введя упрощающее положение о постоянстве температуры. В точности по аналогии с моделью из статьи [6] мы будем обозначать множество всех реакций как \mathcal{R} , где отдельная реакция $\rho \in \mathcal{R}$ имеет вид:



В этом случае $\text{Reaction}_m^{(2)}$ будет выражено ($A(\rho)$ — субстраты, $B(\rho)$ — продукты для ρ) так:

$$\begin{aligned} \text{Reaction}_m^{(2)} &= \sum_{\substack{\rho \in \mathcal{R}: \\ m \in A(\rho)}} k_m^+(\rho) \cdot \omega_{(2)\rho}^- + \sum_{\substack{\rho \in \mathcal{R}: \\ m \in B(\rho)}} k_m^-(\rho) \cdot \omega_{(2)\rho}^+ - \sum_{\substack{\rho \in \mathcal{R}: \\ m \in A(\rho)}} k_m^+(\rho) \cdot \omega_{(2)\rho}^+ - \sum_{\substack{\rho \in \mathcal{R}: \\ m \in B(\rho)}} k_m^-(\rho) \cdot \omega_{(2)\rho}^-, \\ \omega_{(2)\rho}^+ &= v^+(\rho) \cdot e^{\frac{-E^+(\rho)}{R \cdot T}} \cdot \prod_{j=1}^{n(\rho)} \left(\frac{N_{a_j}^{(2)} \cdot V_{a_j}}{V_{nuc}} \right)^{k_{a_j}^+(\rho)} \cdot \min_i \left(\frac{N_{a_i}^{(2)}}{k_{a_i}^+(\rho)} \right), \\ \omega_{(2)\rho}^- &= v^-(\rho) \cdot e^{\frac{-E^-(\rho)}{R \cdot T}} \cdot \prod_{j=1}^{p(\rho)} \left(\frac{N_{b_j}^{(2)} \cdot V_{b_j}}{V_{nuc}} \right)^{k_{b_j}^-(\rho)} \cdot \min_i \left(\frac{N_{b_i}^{(2)}}{k_{b_i}^-(\rho)} \right). \end{aligned}$$

В этих выражениях отдельные факторы имеют следующую интерпретацию:

- 1) $\prod_{j=1}^{n(\rho)} \left(\frac{N_{a_j} \cdot V_{a_j}}{V_{nuc}} \right)^{k_{a_j}^+(\rho)}$ — оценка вероятности столкновения реагентов для реакции ρ ;
- 2) $\exp \frac{-E(\rho)}{R \cdot T}$ — оценка вероятности инициации реакции при столкновении реагентов;
- 3) $\min_i \left(\frac{N_{a_i}}{k_{a_i}^+(\rho)} \right)$ — максимально возможное число реакций ρ в единицу времени;
- 4) $v(\rho)$ — элементарная скорость реакции (при условии, что она уже иницирована).

⁴Пример такого процесса — сплайсинг (вырезание интронов) сразу после транскрипции РНК.

Замечание 6. Использование изложенной модели химической кинетики подразумевает обязательное разложение всех рассматриваемых реакций на элементарные стадии. В противном случае параметры элементарной скорости $v(\rho)$ и энергии активации $E(\rho)$ нельзя будет считать постоянными величинами, а уравнения для баланса реакций потеряют физическую строгость.

Воздействие на активность генов g моделируется через изменение факторов $I_{p^r}(g)$.

$$\begin{aligned} \frac{dI_{p_j^r}(g_i)}{dt} = & \sum_{k: \alpha(\dots, p_k^r) > 0} \alpha(g_i, p_j^r, p_k^r) \sigma \left(I^{\text{lim}}(g_i, p_j^r, p_k^r) - I_{p_j^r}(g_i) \right) \left(\frac{V_{p_k^r} N_{p_k^r}^{(2)}}{V_{nuc}} \right) I_{p_k^r}(g_i) + \\ & + \sum_{k: \alpha(\dots, p_k^r) < 0} \alpha(g_i, p_j^r, p_k^r) \sigma \left(I_{p_j^r}(g_i) - I^{\text{lim}}(g_i, p_j^r, p_k^r) \right) \left(\frac{V_{p_k^r} N_{p_k^r}^{(2)}}{V_{nuc}} \right) I_{p_k^r}(g_i). \end{aligned}$$

В этом выражении $\alpha(g_i, p_j^r, p_k^r)$ — это скорость увеличения (уменьшения) вероятности присоединения p_j^r к гену g_i под действием p_k^r , а $I^{\text{lim}}(g_i, p_j^r, p_k^r)$ — это пороговое значение, до которого будет повышаться (или понижаться) значение $I_{p_j^r}$. В качестве пороговой функции $\sigma(x)$ можно использовать либо функцию Хевисайда, либо приближенную к ней функцию сигмоиду $\sigma_c(x) = 1/(1 + e^{-x/c})$, при $c \rightarrow 0$ и $c > 0$.

Замечание 7. Важной особенностью уравнений для пересчета $I_{p_j^r}(g_i)$ является использование рекурсивной схемы для моделирования регуляторных воздействий. В частности, за счет умножения на $I_{p_k^r}(g_i)$ мы можем моделировать как последовательные воздействия с образованием комплексов из регуляторных молекул на промоторе гена, так и воздействия, приводящие к изменению плотности упаковки ДНК этого гена.

Поскольку для определенных типов генетической регуляции возможно расходование регуляторных агентов при их реализации, то мы вводим для них специальное слагаемое, отвечающее за их потребление (коэффициент β описывает скорость потребления):

$$\begin{aligned} \frac{dN_{p_j^r}^{(2)}}{dt} = & \sum_{\alpha(g_i, p_j^r, p_k^r) > 0} \beta(g_i, p_j^r, p_k^r) \sigma \left(I^{\text{lim}}(g_i, p_j^r, p_k^r) - I_{p_j^r}(g_i) \right) \left(\frac{V_{p_k^r} N_{p_k^r}^{(2)}}{V_{nuc}} \right) I_{p_k^r}(g_i) + \\ & + \text{Reaction}_{p_j^r}^{(2)} + \sum_{\alpha(g_i, p_j^r, p_k^r) < 0} \beta(g_i, p_j^r, p_k^r) \sigma \left(I_{p_j^r}(g_i) - I^{\text{lim}}(g_i, p_j^r, p_k^r) \right) \left(\frac{V_{p_k^r} N_{p_k^r}^{(2)}}{V_{nuc}} \right) I_{p_k^r}(g_i). \end{aligned}$$

Для описания динамики веществ в пределах цитоплазмы мы будем использовать полностью аналогичный подход, за тем лишь исключением, что вместо объема ядра V_{nuc} мы будем использовать объем цитоплазмы V_{ret} , содержащий эндоплазматический ретикулум. В частности, если отдельный белок p непосредственно синтезируется в результате трансляции с молекулы РНК r_i , то данный факт будет отражен в модели следующим образом:

$$\frac{dN_p^{(1)}}{dt} = v_{p_2^r}(r_i) \left(\frac{V_{p_2^r} N_{p_2^r}^{(1)}}{V_{ret}} \right) \left(\frac{V_{r_i} N_{r_i}^{(1)}}{V_{ret}} \right) I_{p_2^r}(r_i) - \text{Exchange}_p^{(1 \leftrightarrow 2)} + \text{Reaction}_p^{(1)}.$$

В этом выражении $v_{p_2^r}$ — это средняя скорость трансляции РНК молекулы r_i на рибосоме p_2^r , а $I_{p_2^r}(r_i)$ — это вероятность присоединения r_i к рибосоме. Если молекула p не может синтезироваться непосредственно в результате трансляции РНК, то тогда в выражении для p будет отсутствовать первое слагаемое. Полностью аналогично записывается и формула для изменения численностей p^r . Отметим, что слагаемое, отвечающее за реакции $\text{Reaction}_p^{(1)}$, составляется с использованием численностей молекул в области цитоплазмы (1), а также объема V_{ret} ретикулума вместо объема V_{nuc} клеточного ядра.

§ 3. Примеры использования модели

В первую очередь мы рассмотрим примеры формализации параметров и переменных модели для таких процессов, как модификация концевых участков гистонов, образование регуляторного комплекса на промоторе гена, а также метилирование ДНК.

Пример 1. Регуляторная молекула p^r укорачивает концевые участки гистонов, что влечет за собой свертывание сегмента ДНК гена g за время τ . В этом случае для всех остальных регуляторных молекул p_j^r мы можем оценить скорость $\alpha(g, p_j^r, p^r) \simeq -1/\tau$. Если молекула p_j^r осуществляет регуляторное воздействие за счет связывания с промотором гена, то для нее порог $I^{\text{lim}}(g, p_j^r, p^r) = 0$. В результате после осуществления воздействия будет заблокирована транскрипция и ни один из факторов транскрипции не сможет присоединиться к промотору гена. Если регуляторное воздействие p_j^r связано с модификацией белков гистонов и не требует прикрепления к промотору гена, то пороговое значение может отличаться от нуля $I^{\text{lim}}(g, p_j^r, p^r) > 0$. Таким образом, после его осуществления ген будет доступен только для факторов ремоделирования хроматина.

Пример 2. Регуляторная молекула p_k^r за время τ осуществляет метилирование CpG-динуклеотидов в промоторной области гена g . В этом случае регуляторное воздействие приведет к увеличению вероятности присоединения факторов конденсации ДНК p^r , рассмотренных в предыдущем примере. Это увеличение будет произведено со средней скоростью $\alpha(g, p^r, p_k^r) \simeq 1/\tau$, а итоговый порог $I^{\text{lim}}(g, p^r, p_k^r) > 0$ можно оценить экспериментально, причем, как правило, в среднем он будет $I^{\text{lim}}(g, p^r, p_k^r) \gtrsim 0.5$.

Пример 3. Регуляторный комплекс из двух молекул p_3^r и p_4^r повышает частоту считывания РНК с гена g в два раза. При этом первым присоединяется молекула p_3^r за время τ_1 , а затем к ней добавляется p_4^r за время τ_2 после присоединения p_3^r . Переведя эти данные на язык нашей модели, получим $\alpha(g, p_4^r, p_3^r) = 1/\tau_1$ и $I^{\text{lim}}(g, p_4^r, p_3^r) = 1$. При этом подразумевается, что начальное значение $I_{p_4^r}(g, 0) = 0$ в нулевой момент времени. Измеренную экспериментально (до регуляторного воздействия) начальную частоту присоединения РНК-полимеразы p_1^r к гену g мы обозначим как $I_{p_1^r}(g, 0) \leq 0.5$. Тогда итоговое регуляторное воздействие будет формализовано как $\alpha(g, p_1^r, p_4^r) = 1/\tau_2$ и $I^{\text{lim}}(g, p_1^r, p_4^r) = 2I_{p_1^r}(g, 0)$.

Во вторую очередь обратимся к примерам использования нашей биохимической модели для уточнения параметров моделей⁵ морфогенеза. Как правило, модели морфогенеза описывают динамику скоплений клеток на основе дифференциальных уравнений, либо применяют имитационный подход (примеры см. в [7, 8]). Общим элементом для этих моделей является использование таких терминов, как типы клеток, пролиферация, дифференцировка, апоптоз, сигнальный обмен между клетками. Мы покажем на примерах, как эти понятия могут быть проинтерпретированы, а отвечающие за них коэффициенты уточнены с помощью биохимической модели клетки.

Пример 4. Под термином *тип клетки* подразумевается текущая специализация клетки и какие у нее могут быть потенциальные потомки в результате деления. Оценить, к какому типу относится клетка на основе биохимической модели, можно в первую очередь на основе концентраций белков p и белковых комплексов p^r внутри этой клетки. Определить тип клетки на практике можно с помощью алгоритмов кластеризации, если заданы эталонные значения концентраций белков для всех типов клеток.

Пример 5. Под *межклеточным сигнальным обменом* подразумевается процесс передачи сообщений между различными клетками скопления, который в обязательном порядке завершает формирование (или высвобождение) вторичных мессенджеров внутри принимающей его

⁵Моделей, которые описывают образование сложных форм из клеточных скоплений.

клетки. При этом передача сообщения на межклеточном пространстве может проводиться различными путями: диффузией специальных молекул-морфогенов, созданием сигнальных контактов между соседними клетками и даже за счет передачи механических напряжений в ткани. Единственное, что должно выполняться для межклеточного сигнального обмена, — это наличие четкого механизма генерации, передачи и приема сигналов клетками. При этом прием обязан завершаться формированием молекул p^r внутри клетки, которые переведут послание на язык генетической регуляции. В рамках модели мы можем описать последний этап такого воздействия за счет изменения концентраций соответствующих p^r внутри клетки и провести сравнительное моделирование: с внешним воздействием и без него.

Пример 6. Под *апоптозом* понимается программируемая клеточная смерть. Этот процесс запускается с помощью специальных факторов апоптоза, чьи концентрации p и p^r будут служить маркерами для регистрации начала данного события в нашей модели. В частности, запуск апоптоза может быть вызван как итоговый ответ на некоторое внешнее сигнальное сообщение, рассмотренное в предыдущем примере.

Пример 7. Под термином *пролиферация* подразумевается рост ткани за счет деления клеток. На уровне отдельной клетки этот рост регулируется контрольными точками клеточного цикла, которые мы можем описать в нашей модели через концентрации соответствующих p и p^r . Очевидно, что изменение скорости прохождения этих точек также может быть вызвано в результате действия сигналов от других клеток. В результате, получив численную оценку того, насколько быстрее или медленнее стала клетка проходить эти контрольные точки, можно получить оценку для параметра ускорения пролиферации $K(\tau_1, \tau_2)$ клеток типа τ_1 под действием сигналов от τ_2 .

Пример 8. Под *дифференцировкой* понимается выбор новых типов для дочерних клеток после деления материнской. Этот выбор обуславливается множеством факторов, но для нас важнейшими будут концентрации белков на момент начала деления клетки, а также наличие метилированных генов в ядре. Так же как и в предыдущих случаях, изменение дифференцировки возможно под действием внешних сигналов. В этом случае мы можем описать это в простейшем варианте с помощью параметра $D(\tau_1, \tau_2)$ — тип клетки, который должны получить клетки τ_1 под действием сигналов от τ_2 .

Заключение

Наиболее важной особенностью нашей модели является ее ориентация на точное и корректное описание кинетики процессов регуляции биосинтеза белков. Этот подход позволяет не только использовать нашу модель как разовый инструмент для установления характера и типа связей между генами и их продуктами, но и в перспективе задействовать ее при иерархическом построении общей модели биохимии клетки (примеры см. в [9]). В контексте данного вопроса следует обсудить насколько в принципе наша базовая модель удовлетворяет четырем требованиям из введения.

Первое требование об *учете промежуточных молекул в регуляторных воздействиях* было удовлетворено частично. Фактически эти молекулы учитываются в механизмах контроля транскрипции, посттранскрипционной обработке и посттрансляционной обработке. При построении обобщения можно будет учесть также и другие молекулы, задействованные в реализации трансляции, ядерного транспорта и трансляции.

Второе, третье и четвертое требования были проработаны существенно лучше: уравнения модели используют строгие физические основания, и все параметры модели имеют биологический смысл и могут быть потенциально измерены экспериментально. Для достижения этой цели мы разделили переменные на два блока: численности в ядре и численности в цитоплазме, а также задействовали формулу для вероятности сложного события и применили принцип геометрической вероятности при выводе уравнений модели. При этом среди параметров модели следует отдельно отметить следующие: $v_{p_1^r}(g_i)$ — скорость синтеза РНК на основе ДНК

гена g_i и $v_{p_2^r}(r_i)$ — скорость синтеза белка на базе РНК r_i . Параметр $v_{p_1^r}(g_i)$ можно оценить, зная, какие нуклеотиды задействованы в последовательности g_i : сколько их в сумме и каковы средние концентрации этих нуклеотидов в ядре. Точно так же параметр $v_{p_2^r}(r_i)$ оценивается на основании того, какие аминокислоты будут задействованы в итоговом белке: сколько их в общей цепочке, и каковы концентрации тРНК с этими аминокислотами в цитоплазме. Оба этих параметра полагаются константами ввиду гипотезы о насыщении ядра достаточным числом нуклеотидов, а цитоплазмы — тРНК с аминокислотами для параллельного синтеза этих молекул. При обобщении модели можно перейти к более точной схеме, в которой данные скорости будут функциями от соответствующих концентраций (нуклеотидов и тРНК).

Важной особенностью нашей модели является использование элементарной подмодели химической кинетики из работы [6] для описания реакций в ядре и цитоплазме: $\text{Reaction}_m^{(2)}$ и $\text{Reaction}_m^{(1)}$. Данный выбор не является принципиальным, и при необходимости вполне можно заменить эту химическую подмодель на другую. В частности, модель химической кинетики из работы [6] требует разложения всех реакций на элементарные стадии. Это вполне может стать существенным ограничивающим фактором, если требуется учитывать в модели реакции с десятками элементарных стадий и более. В этом случае разумнее будет использовать упрощенные схемы для описания химических реакций. Примером такой упрощенной схемы могут послужить наши схемы, задействованные при описании транскрипции РНК и при описании трансляции белков.

СПИСОК ЛИТЕРАТУРЫ

1. Alberts B., Johnson A., Lewis J., Morgan D., Raff M., Roberts K., Walter P. Molecular biology of the cell. 6th edition. New York: Garland Science, 2014.
2. Karlebach G., Shamir R. Modelling and analysis of gene regulatory networks // Nature Reviews Molecular Cell Biology. 2014. Vol. 9. P. 770–780. DOI: [10.1038/nrm2503](https://doi.org/10.1038/nrm2503)
3. Hecker M., Lambeck S., Toepfer S., van Someren E., Guthke R. Gene regulatory network inference: Data integration in dynamic models — a review // Biosystems. 2009. Vol. 96. Issue 1. P. 86–103. DOI: [10.1016/j.biosystems.2008.12.004](https://doi.org/10.1016/j.biosystems.2008.12.004)
4. Колчанов Н.А., Игнатъева Е.В., Подколотная О.А., Лихошвай В.А., Матушкин Ю.Г. Генные сети // Вавиловский журнал генетики и селекции. 2013. Вып. 17. № 4/2. С. 833–850.
5. Davidson E., Levin M. Gene regulatory networks // Proceedings of the National Academy of Sciences of the United States of America. 2005. Vol. 102. No. 14. P. 4935. DOI: [10.1073/pnas.0502024102](https://doi.org/10.1073/pnas.0502024102)
6. Nazarov M.N. New approaches to the analysis of the elementary reactions kinetics // Journal of Siberian Federal University. Mathematics & Physics. 2014. Vol. 7. Issue 3. P. 373–382.
7. Назаров М.Н. Моделирование роста ткани с учетом возможности внешнего воздействия на ее форму // Прикладная дискретная математика. 2013. № 4. С. 103–113.
8. Назаров М.Н. О поиске универсальных параметров для моделей морфогенеза // Владикавказский математический журнал. 2013. Т. 15. № 3. С. 58–66.
9. Castellanos M., Wilson D.B., Shuler M.L. A modular minimal cell model: Purine and pyrimidine transport and metabolism // Proceedings of the National Academy of Sciences of the United States of America. 2004. Vol. 101. No. 17. P. 6681–6686. DOI: [10.1073/pnas.0400962101](https://doi.org/10.1073/pnas.0400962101)

Поступила в редакцию 09.08.2016

Назаров Максим Николаевич, старший преподаватель, кафедра высшей математики 1, Национальный исследовательский университет «МИЭТ», 124498, Россия, г. Москва, г. Зеленоград, площадь Шокина, 1. E-mail: nazarov-maximilian@yandex.ru

M. N. Nazarov

The basic mathematical model for the description of regulatory processes of protein biosynthesis

Citation: *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, 2016, vol. 26, issue 4, pp. 515–524 (in Russian).

Keywords: intercellular biochemistry, gene regulatory networks, gene expression regulation, nuclear transport.

MSC2010: 92C40, 92C42, 80A30

DOI: [10.20537/vm160406](https://doi.org/10.20537/vm160406)

In this paper we propose a model to describe the kinetics of the main processes regulating protein biosynthesis. For the resulting model we include above all the following processes: regulation of RNA transcription, splicing, nuclear transport, as well as regulation of the RNA translation. The main objective of this model is to provide some basic tools for the modelling of the intracellular biochemistry, which will be based on physically rigorous equations and biologically sound coefficients. To achieve this goal, we divide variables into two blocks: the number in the nucleus and the number in the cell cytoplasm for proteins and RNA, and apply the principle of geometric probability. Moreover, all the effects on gene transcription are modelled by using special variables, the probabilities of attachment of the regulatory complex to the gene. A specific example of such a complex is the RNA polymerase, and accordingly, by changing the probability of its connection to DNA one will change directly the rate of the gene transcription initiation.

REFERENCES

1. Alberts B., Johnson A., Lewis J., Morgan D., Raff M., Roberts K., Walter P. *Molecular biology of the cell, 6th edition*, New York: Garland Science, 2014.
2. Karlebach G., Shamir R. Modelling and analysis of gene regulatory networks, *Nature Reviews Molecular Cell Biology*, 2014, vol. 9, pp. 770–780. DOI: [10.1038/nrm2503](https://doi.org/10.1038/nrm2503)
3. Hecker M., Lambeck S., Toepfer S., van Someren E., Guthke R. Gene regulatory network inference: Data integration in dynamic models — a review, *Biosystems*, 2009, vol. 96, issue 1, pp. 86–103. DOI: [10.1016/j.biosystems.2008.12.004](https://doi.org/10.1016/j.biosystems.2008.12.004)
4. Kolchanov N.A., Ignat'eva E.V., Podkolodnaya O.A., Likhoshvai V.A., Matushkin Yu.G. Gene nets, *Vavilov Journal of Genetics and Breeding*, 2013, vol. 17, no. 4/2, pp. 833–850 (in Russian).
5. Davidson E., Levin M. Gene regulatory networks, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, vol. 102, no. 14, p. 4935. DOI: [10.1073/pnas.0502024102](https://doi.org/10.1073/pnas.0502024102)
6. Nazarov M.N. New approaches to the analysis of the elementary reactions kinetics, *Journal of Siberian Federal University. Mathematics & Physics*, 2014, vol. 7, issue 3, pp. 373–382.
7. Nazarov M.N. Modelling the tissue growth with the possibility of external influence on tissue shape, *Prikladnaya Diskretnaya Matematika*, 2013, no. 4, pp. 103–113 (in Russian).
8. Nazarov M.N. The search for universal parameters for the models of morphogenesis, *Vladikavkaz. Mat. Zh.*, 2013, vol. 15, no. 3, pp. 58–66 (in Russian).
9. Castellanos M., Wilson D.B., Shuler M.L. A modular minimal cell model: Purine and pyrimidine transport and metabolism, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, vol. 101, no. 17, pp. 6681–6686. DOI: [10.1073/pnas.0400962101](https://doi.org/10.1073/pnas.0400962101)

Received 09.08.2016

Nazarov Maksim Nikolaevich, Senior Lecturer, Department of Higher Mathematics 1, National Research University of Electronic Technology, pl. Shokina, 1, Zelenograd, Moscow, 124498, Russia.

E-mail: nazarov-maximilian@yandex.ru